DOCUMENT RESUME

ED 232 864 SE 042 638

AUTHOR	Ashcraft, Mark H.
TITLE	Simulating Network Retrieval of Arithmetic Facts.
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
REPORT NO	LRDC-1983/10
PUB DATE	83
NOTE	59p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Addition; Adult Learning; *Arithmetic; Cognitive Development; Cognitive Processes; Decision Making; Long Term Memory; *Mathematics Education; *Models; Multiplication; *Reaction Time; *Recall (Psychology); Schemata (Cognition); *Simulation
IDENTIFIERS	Mathematics Education Research; *Mental Computation; National Science Foundation

ABSTRACT

This report describes a simulation of adults' retrieval of arithmetic facts from a network-based memory representation. The goals of the simulation project are to: demonstrate in specific form the nature of a spreading activation model of mental arithmetic; account for three important reaction time effects observed in laboratory investigations; and provide a basis from which developmental changes in performance can be meaningfully understood. The simulation, based on the process of spreading activation among related nodes, retrieves correct answers as the activation from a problem's addends/multipliers intersects in the network. Two decision processes are simulated. The first consists of an attempted matching of retrieved and stated answers, essentially a process of number inequality judgments. The second decision mechanism involves competition among related nodes, based on levels of activation generated during the search phase for both retrieved and stated answers. Both decision mechanisms are necessary to predict the separate effects of split and confusions which have been observed in the laboratory. The conclusion of the paper discusses the relationship of the simulation to developmental processes in mental arithmetic performance, considering the interplay between declarative knowledge of the sort simulated and "procedural" knowledge of arithmetic algorithms and heuristics. (Author/JN)

SIMULATING NETWORK RETRIEVAL OF ARITHMETIC FACTS

Mark H. Ashcraft

Cleveland State University

Learning Research and Development Center

University of Pittsburgh

1983

The research reported herein was supported by the Learning Research and Development Center, supported in part as a research and development center by funds from the National Institute of Education (NIE), Department of Education. The opinions expressed do not necessarily reflect the position or policy of NIE and no official endorsement should be inferred.

Simulating Network Retrieval of Arithmetic Facts

Mark H. Ashcraft

Cleveland State University

Title Footnote

The project described here was largely accomplished during my sabbatical leave, when I was a Visiting Scholar at the Learning Research and Development Center, at the University of Pittsburgh. The work was generously supported by the LRDC, which is supported in part by funds from the National Institute of Education. The research upon which the model is based was partially supported by a National Science Foundation grant, SED-8021521, through the Research in Science Education (RISE) program; this grant also provided support for the final revisions to the model. I wish to thank Dr. James Greeno for making the LRDC facilities available to me during my sabbatical leave, and for the many discussions we had during the development of this simulation.

Send proofs to:

Mark H. Ashcraft

Department of Psychology

Cleveland State University

Cleveland, OH 44115

Running Head: Retrieval of Arithmetic Facts

Abstract

This paper describes a simulation of adults' retrieval of arithmetic facts from a network-based memory representation. The simulation is based on the process of spreading activation among related nodes, and retrieves correct answers as the activation from a problem's addends/multipliers intersects in the network. Two decision processes are simulated. The first consists of an attempted matching of retrieved and stated answers, essentially a process of number inequality judgments. The second decision mechanism involves competition among related nodes, based on levels of activation generated during the search phase for both retrieved and stated answers. Both decision mechanisms are necessary to predict the separate effects of split and confusions which have been observed in the laboratory. The conclusion of the paper discusses the relationship of the simulation to developmental processes in mental arithmetic performance.

This report describes a simulation model of memory retrieval, specifically the retrieval from organized memory of simple addition and multiplication facts. The empirical results modeled by this simulation are found in several recent publications (e.g., Ashcraft & Stazyk, 1981; Stazyk, Ashcraft, & Hamann, 1982) which examined the chronometric aspects of adults' performance. A number of developmental investigations of performance on the "basic arithmetic facts" have also been conducted (e.g., Ashcraft & Fierman, 1982), and the final section of this report considers the developmental implications of the model. The most basic assumptions of the model are, first, that knowledge of arithmetic facts, such as $4 + 3 = 7$ and $5 \times 6 = 30$, is stored in interrelated network representations in long term memory, and second that a process of spreading activation through these networks is the mechanism for accessing this information (e.g., Anderson, 1983; Collins & Loftus, 1975). The three goals of the simulation project are: first, to demonstrate in specific form the nature of a spreading activation model of mental arithmetic performance; second, to account for three important reaction time (RT) effects observed in laboratory investigations; and third, to provide a basis from which developmental changes in performance can be meaningfully understood.

This report is organized into four major sections. First, a brief review is provided of the three most important empirical effects in this area, the effects of problem size, split or symbolic distance, and confusions. These three effects, robust empirical phenomena demonstrated in both adult and child populations, are viewed as critical results for any serious model of mental arithmetic. Following a description of the model, an account of its predictions across relevant experimental conditions is presented. The success of the model both in performing correctly on the relevant tasks and in predicting the three critical RT effects is viewed as relatively strong support for the general search and decision framework proposed by Ashcraft (1982). The lack of fit between the model and the data are considered from the standpoint of the adequacy of the theory of mental arithmetic, with attention to the possibility that the departures reflect non-retrieval processing of arithmetic facts. The final section of the report discusses some possible directions of development for the simulation of mental arithmetic performance, especially those aspects of the model which will need attention for a more complete understanding of the development of mental arithmetic processes. Specifically, the interplay between declarative knowledge of the sort simulated here and "procedural" knowledge of arithmetic algorithms and heuristics is considered in this final section.

Empirical Effects in Studies of Mental Arithmetic

Across a period of several years, we have been investigating the chronometric effects obtained when children and adults are presented with simple arithmetic facts in a laboratory situation. Our studies, as well as those of other investigators (e.g., Svenson, 1975), have shown a great consistency in major findings, and a theoretical proposal for understanding these effects has been described in a recent review paper (Ashcraft, 1982). The three results of particular significance that I focus on here are the dependency of RT on the size of the problem being processed, the variation of RT as a function of degree of incorrectness or "split" in false problems, and the intriguing relationships found in RTs under conditions of addition and multiplication "confusions". These effects are described briefly here, along with an account of their relationship to the network-decision model presented in Ashcraft (1982).

Problem Size Effect

The most fundamental empirical result in the study of mental arithmetic performance is the problem size effect. Stated simply, the time to process a simple arithmetic problem, either in addition or multiplication, increases as the size of the problem's answer increases. This problem size effect is obtained regardless of the age group under study, although its parameters vary developmentally. The effect emerges in both laboratory tasks which characterize this area of research, the production task, which requires generating the answer to a problem (e.g., 4 + 3 = ?), and the verification task, which requires judging a stated answer as true or false (e.g., 4 + 3 = 9).

It is widely agreed that the problem size effect derives from a mental search or compute operation. Our empirical results (see the extensive discussion in Ashcraft, 1982) have indicated that counting is not a particularly common solution method to adults when processing the basic arithmetic facts (Ashcraft & Stazyk, 1981; Stazyk et al., 1982). As a consequence, we have proposed a model of adult processing based on fact retrieval from organized memory. In this model, the problem size effect is attributed to the search stage of processing. During this stage, retrieval is assumed to operate on a memory representation of addition (and multiplication) facts, a representation organized as an interrelated network of nodes. Upon presentation of a simple problem, say 4 + 5 = 9, the model claims that a directed search through memory is initiated, activation spreading from the "parent" nodes of 4 and 5 to the "families" linked to those nodes, the "4 + " facts and the " + 5" facts. When an intersection of these two sources of activation occurs, then the correct answer has been located in the network. Importantly, the time to achieve an intersection is assumed to vary directly as a function of "semantic distance", the distance traversed through the network from the origin to the intersection node. While various structural variables (like the sum or the correct product) provide an index of this search distance, Stazyk et al. pointed out that equally predictive results are obtained when normatively derived measures of problem difficulty are used.

Accordingly, the problem size effect in this simulation is a reflection of semantic distance traversed during an intersection search in the network of stored facts. This intersection search process is assumed to be operative in all retrievals of simple addition and multiplication facts; that is, the somewhat larger addition problems studied in Ashcraft and Stazyk's Experiment II revealed similar evidence of the distance effect in search, where the size of the unit's column addition contributed to the prediction of overall RT (for example, the embedded 4 + 3 = 7 in the problem 14 + 13 = 27). Thus, the principle of the simulation, spreading activation, is conceptualized as common to not only the basic addition and multiplication facts but also to more complex problems using the facts as components. Secondly, the empirical evidence suggests that search processes as conceived here are general to both the verification task (true/false judgments) as well as the production task. No important differences, other than the absence of a true/false decision stage in production, have been isolated in our studies of this task factor (Ashcraft & Bartolotta, Note 1; Fierman, 1980).

Symbolic Distance/Split Effect in Decision

On the presumption that the search stage of processing passes the correct answer to the decision mechanism, a same/different judgment must occur during decision in order for the subject to respond correctly. Somewhat surprisingly, this decision process does not reveal a simple yes/no RT function, but instead a function which depends on the magnitude of difference between the retrieved answer and the answer stated in the stimulus (Moyer & Landauer, 1967). In the Ashcraft and Battaglia (1978) report, problems wrong by 1 or 2 (verification task) were approximately 130 msec slower than their true counterparts, while those wrong by 5 or 6 were only 30 msec slower. Ashcraft and Stazyk examined splits of 1, 5, 9, and 13, and found comparable RT effects at the three lower values of split. In its basic form, then, the split effect reveals a decrease in decision time as the numerical difference between correct and incorrect answers increases.

According to several theorists (see Banks, 1977, for a review), this effect is one of discriminability between symbolic (number) stimuli, such that the nearer two stimuli are on some mental continuum of magnitude, the more difficult the resultant discrimination between them. Banks, Fujii, and Kayra-Stuart (1976), Friedman (1978), Holyoak and Walker (1976) and others have documented this inverse relationship between symbolic distance and decision time in tasks involving comparisons on dimensions of time, temperature, size, and even evaluative dimensions lacking any obvious numerical referents. Thus, the split or symbolic distance effect is more indirect than some process of computing differences between correct and stated answers; indeed, a subtraction-like basis for the effect is totally implausible in tasks like those used by Holyoak and Walker (decide "Which is hotter — lukewarm or torrid"). This effect of split, furthermore, applies to multiplication performance as well as addition (Stazyk et al., 1982), and reveals a proportional basis for the effect (that is, the proportion by which the split departs from the correct answer seems to be the operative factor, rather than simply the numerical difference). The split effect has never failed to emerge quite strongly in our experiments, and in fact is especially pronounced in children of early grade school age. The simulation to be presented, clearly, must predict this symbolic distance effect in performance.

Effects of Confusions

In a little noticed paper, Winkelman and Schmidt (1974) found evidence of significant interrelationships in memory between addition and multiplication knowledge. In their report, some distractor stimuli were given answers correct under the alternate arithmetic operation, such as 4 + 5 = 20 and 6 X 3 = 9. Their results showed a marked slowing of RT as a function of these "associative confusions", a slowing taken as evidence for associations between the two operations in memory (or more precisely, competing associations between pairs of numbers, say 4 and 5, where the association to 9 through addition competes with the other to 20 through multiplication). In our own work, a comparable confusion effect was obtained in a study of multiplication (which avoided the cross-operation manipulation in Winkelman and Schmidt's procedures).

Specifically, Stazyk et al. found that multiplication problems given an incorrect answer which was a multiple of one of the problem's numbers (e.g., 7 X 8 = 49) were up to 300 msec slower than distractors which did not use a multiple as the incorrect answer. This confusion effect obtained whether the stated answer was too small or too large, thus ruling out a counting-up explanation of the effect. Duffy and Fisher (Note 2) have reported the same confusion effect, as well as significant slowing of RT when a legal vs. illegal (e.g., 35 vs. 33) false answer is presented. In short, significant interrelationships among multiplication facts slow down the retrieval of this knowledge when competing information is activated.

In the absence of more recent data (but see Bartolotta, Note 3), both the within-operation and cross-operation confusion effects are viewed as important to an adequate theory. This importance goes beyond the mere requirement that a model should predict the observed effects, however. In particular, the assumption that interrelated information is highly similar and hence confusable is a hallmark of network approaches to memory representations. The emergence of such relatedness or "confusion" effects in empirical studies constitutes compelling evidence <u>for</u> a network representation of arithmetic facts. Consequently, the present simulation embodies relatedness among nodes as a fundamental structural assumption.

The Simulation Model

Figure 1 illustrates the overall theoretical approach to mental arithmetic performance taken in the simulation. The figure is from Ashcraft's review of chronometric evidence in the development of mental arithmetic processes. The descriptive model in the figure shows a standard four stage information processing model of performance, with sequential arrangement of encoding, search, decision, and response stages. As illustrated, the stimulus problem is first encoded, then passed to a search/compute stage. Time for completion of stage two is predicted to be a ioint function of memory retrieval time and "procedural knowledge" involvement. That is, as the subject's verification or production performance depends more heavily on aspects of procedural knowledge like counting, estimating, or solving by a rule, the reaction time increases.

Figure 1 about here

Search time through the declarative knowledge structure is assumed to be more rapid than solving via some procedural basis like counting, and further is viewed as a reflection of semantic distance in the network of stored facts [1] Activation originating from the problem's numbers spreads through the network of stored facts, with retrieval dependent on finding an intersection. Once the search stage is completed, output to the decision mechanism includes the correct answer, as retrieved from the network, and information concerning the levels of activation generated by the search. Upon making a decision as to the correctness of the stated answer (assuming a verification task), control is passed to the response stage for preparation for the relevant response. As is standard in such models, overall RT is assumed to be the additive combination of times for each of the four stages.

The major activities of interest in this illustration, and in the simulation, are those associated with the intermediate search/compute and decision stages. No particularly relevant information exists on the encoding and response phases of performance in these tasks, so the simulation is silent on these processes. Figure 2 presents a summary flow chart of processing activities in the simulation, with RT prediction equations at the bottom.

Figure 2 about here

Network Structure

As discussed above, the governing assumption embedded in the simulation is the notion of network storage of addition and multiplication facts in a semantic-like interrelated structure. In the model, two "mental tables" exist, one for addition and one for multiplication. Each table is a 10 X 10 matrix of entries, with column and row headings of the digits zero through nine. These headings, hereafter referred to as "parent nodes", are the entry points of the search process. Each parent points to the entries stored along its column or row, referred to as the parent's "family" of facts. Thus, the digit 4, when it appears as the first addend (or multiplier) in a problem, corresponds to the parent node of the "4 + b" family of entries (or the "4 X b" family in multiplication). The family corresponds to the 10 combinations 4 + b where b=(0,1,2...9). Likewise, if the second addend (multiplier) in a problem is 7, the parent node for the "+ 7" facts points to the ten entries for "a + 7", where a=(0,1,2...9).

Each table in this scheme consists of 100 nodes, the 100 possible intersections of the two parents a and b. Each node contains two pieces of information, the first of which is the correct answer to the particular combination of a and b under the appropriate operation (addition or multiplication). It seems entirely appropriate to store all of the correct answers as part of the model's initial configuration, given the virtually perfect performance evidenced by adults on simple addition and multiplication tasks, whether speeded or not. Such presetting of answers is a questionable procedure in developmental contexts, of course; see the discussion in the final section.

The second piece of information stored in each node is a distance value, representing the accessibility of the node in question. In a very direct way, this feature of the model corresponds to the significant predictor variables selected in regression analyses of the problem size effect on RT. Our early research indicated that the square of the correct sum of an addition problem provided the best index of distance (Ashcraft & Battaglia, 1978), as did correct product in multiplication problems (Stazyk et al., 1982). As Stazyk et al. point out, however, normatively derived measures of distance are equally useful in predicting empirical RT functions. Moreover, a more plausible theoretical rationale can be stated for using difficulty ratings, as opposed to structural variables like correct sum or product. As in like situations which used normatively-based estimates of semantic distance (e.g., Ashcraft, 1976; Glass, Holyoak, & O'Dell, 1974; Loftus & Suppes, 1972; Whaley, 1978), it is assumed here that subjects' ratings are relatively straightforward indices of strength in memory (but see the final discussion for a significant exception). Such strength, further, is widely assumed to vary with important variables like frequency of exposure, age of acquisition, and practice. While the ultimate reason for accessibility values is not elucidated by this assumption, of course, it does suggest certain testable empirical effects. For instance, problems of higher accessibility are more likely to have been encountered frequently during instruction, or to have been "invented" prior to instruction; accessibility values should be altered by extensive practice, or for that matter, extensive neglect.

The particular values used in the present model were derived from subjective ratings of difficulty of the basic 100 facts in addition and multiplication (Ashcraft & Hamann, Note 4; see also Resnick & Ford, 1982, for citations to similar studies). These difficulty ratings, gathered on a 1 to 9 scale, have been converted to "accessibility" scores (expressed as percentages), with the least difficult problems having the highest accessibility. As a consequence, distances are superimposed on the 10 X 10 answer matrices, functionally yielding irregularly shaped table representations. That is to say, entering the usual <u>printed</u> addition table at its 0,0 origin, the "city block" distance to the 6 + 8 = 14 node, for example, is exactly twice the distance to the 4 + 3 = 7 node (or any node with an answer equal to 7). Literal table models of mental representation involve the same distance scheme. In the mental tables simulated here, however, the distance to any particular node is not directly proportional to the city block location of the node, but instead is directly proportional to the accessibility of the node. Figure 3 illustrates the simulation's coding of the answer-accessibility values for the problem 4 + 3 = 7, along with a depiction of the relevant nodes in the multiplication table for the twin problem 4 X 3 = 7.

Figure 3 about here

Two important and related ramifications of this distance scheme require elaboration here, since the non-linear (non-proportional) shape of the mental table is the most easily misinterpreted aspect of the model. Ashcraft and Battaglia referred to a mental table which was "stretched" in the region of larger sums. This stretching was a post hoc modification of the literal square addition table which was suggested by the problem size effect we obtained, RT proportional to the square of the problem's sum. Examination of the table of distances used here suggests that this non-linear effect derives from increasing difficulty of the larger problems; the "stretching" is nicely captured by subjects' ratings on the problems. Furthermore, the distance-based representation of problems accounts for the "tie" effect found repeatedly in the RT data. In a literal-distance mental table, a tie problem like 9 + 9 = 18 requires a lengthy search according to the city block distance scheme, a search distance proportional to 18. Since tie problems generally require only a fairly constant and quite short amount of time in the RT studies, this effect is usually taken as evidence against mental table approaches. In the present simulation, however, distances for tie problems are quite short, reflecting subjects' judgments that these are relatively easy problems. No special tailoring of the tables is necessary for tie problems, in other words, to generate predictions of short search distances (and rapid RTs) -- the effect follows naturally from the accessibility values.

The second important ramification of the present distance-based approach involves relationships among problems. Structurally, both in literal table approaches and in the present simulation, a problem like 4 + 4 = 8 is adjacent to 4 + 3 = 7, by virtue of its status as the successor to 4 + 3 in the "4 + " family. ("Next" relations, pointing forwards and backwards along columns and rows, are assumed throughout the networks. These denote pathways along which the activation spread takes place.) While such adjacency indicates a strong interrelationship between problems, the functional degree of interrelatedness during processing depends not only on structural proximity but also on accessibility. A concrete example will be useful to illustrate this point. The accessibility value for the tie problem 4 + 4 = 8 is 85 (out of 100). Two nodes away, for 4 + 2 = 6, accessibility is 84; two nodes in the other direction, accessibility is 76 for 4 + 6. Thus, while the two problems are equidistant from 4 + 4 in the structural sense, their functional relatedness to 4 + 4 is not symmetrical since they differ in accessibility from the parent nodes. During search, the fact that 4 + 6 and 4 + 2 are two nodes away from the stimulus problem 4 + 4 will be modulated by the differential accessibility of the problems. Thus, the structural relationships encoded in the networks, both "next" and the location equivalences of the addition and multiplication forms of a problem (the answer to 5 + 8 occupys the same location in the addition network as the 40 does in the multiplication network; this relation is termed "other operation"), are modified by the relevant distances associated with the nodes. This feature of the model results in quite specific predictions about inter-problem priming, especially within the multiplication operation.

Search Processes

At the conclusion of the encoding phase, the simulation accesses the procedural knowledge component to begin central processing of the stimulus. At present, this component consists only of a call to the search/retrieval process, resulting in initiation of network retrieval. This does not imply that adults' knowledge about arithmetic is believed to have only this search procedure available. Instead, this reflects the overall intention of the simulation to explore the power of a network-based retrieval model as an explanation of arithmetic performance. To foreshadow one of the modifications of the model offered below, there are data which suggest that network retrieval is not the process by which adults perform multiplication problems of the form a X 0 = 0. Similar arguments might be drawn to force other "special cases" in arithmetic to be computed via some small set of rules (for instance, any problem involving a zero or one as an addend/multiplier; see Ashcraft, in press, and Baroody, in press). While the model illustrated in Figure 1 intended such rules to be available for these cases, the unelaborated simulation is intentionally lacking in those processes. By this method, the range of applicability of the network search scheme may be determined more directly, to be followed by modifications for special cases and other exceptions.

At a general level, the operation of the search stage consists of parallel searches through the networks as a function of the numbers stated in the problem. Activation is passed from each of the three sources a, b, and c (in the problem a +/X b = c), in amounts dictated by the accessibilities of the encountered nodes. That node receiving the highest amount of activation during this search, termed _x_ here, represents the strongest intersection of activations, and is taken by the simulation as the correct answer to the stimulus problem. This node propogates a decreasing gradient of activation to its neighbors, those nodes along both families of facts denoted by the "next" relation, with amounts of activation dependent on structural proximity and individual node accessibility. A comparable gradient is also generated by the "C-nodes" which were activated[2], that is, all nodes with answer values equal to c. Overall then, the networks accumulate activation from the three separate sources a, b, and c. The search stage is completed when the scan through the networks identifies the most highly activated node. The answer and accessibility values stored at that node represent the correct answer to the stimulus problem and the search distance to that problem, respectively.

The details of this spreading activation process illustrate the interplay of distance and degree of relatedness in the search stage as it is conceived here (the simulation program executes sequentially, of course, although the net effect of processing is a simulation of parallel search). Activation is passed from the first and then second parent nodes in the problem to each family member, and then from each C-node in the networks as well. Note that this spread occurs in both the addition and multiplication networks, although at only half the normal amount in the irrelevant operation (as determined by the operator sign + or X in the stimulus). The C-nodes transmit only half the amount of activation to their neighbors that parent nodes do, since the answer in a verification problem is the uncertain and possibly incorrect element of the stimulus. A decreasing gradient of activation is generated around the most highly activated family member, and around each C-node as well, simulating the activation of related nodes in the network structure. Assuming that some node has been activated to at least the criterion level of 100, retrieval of the corect answer has been accomplished. (Note that at least two sources of activation are necessary for a node to pass criterion. Since many nodes will be doubly activated in this fashion, the important "node of intersection" is a convenient short-hand term for that one node, in the full set of nodes which has exceeded criterion, that has the highest activation level. Not surprisingly, the full set of intersection nodes becomes important during the decision stage.)

A concrete example of these processes may clarify the activity simulated during search here. Referring to the values in Figure 3, the processing of the problem 4 + 3 = 7 results in activation of all ten family members along the "4 +" row, the ten nodes along the "+ 3" column, the corresponding 20 nodes in the multiplication network, and the eight nodes (six in addition, two in multiplication) having 7 stored as the answer. The activation levels generated by parents, for convenience, equal the accessibility values associated with each node (e.g., 84 for 4 + 2 = 6, 87 for 1 + 3 = 4). Since activation is summative, the fact that the 4 + 3 = 7 node is twice activated by the spreading activation from parent nodes means that its accessibility value of 81 is doubled, yielding 162. Finally, each C-node receives half the amount of activation denoted by its accessibility value, this activation also summing with any generated from the parent nodes. Concretely then, the target node for 4 + 3 = 7 will have achieved an activation level of approximately 202, the sum of 81 + 81 + half of 81. Assuming that this is the greatest amount of activation in the networks, this node then activates its neighbors through the "next" relations in a decreasing fashion, nearest nodes receiving a larger proportion than further nodes. Specifically, nodes which are immediately adjacent to the highest node of intersection receive 90% of their previous values; those two nodes away receive 80%, and so forth. Exactly the same sort of gradient is applied to the nodes adjacent to the activated C-nodes, with the same summation rule applying throughout. By this means, the activation levels in the networks at the completion of search represent both the inherent accessibilities of the nodes, as measured normatively, and the structural relatedness of the various problems to the retrieved target node, as indexed by proximity. Relationships to nodes in the alternate arithmetic operation (the multiplication network when an addition problem was presented, and vice versa) are captured by the parallel activation of both networks. Thus, when 4 + 3 = 7 is presented, the most highly activated node in the multiplication network is the C-node 12 for 4 X 3 = 12.

The search processes described here simulate a spreading activation search in two principle ways. First, the answer to a problem is indeed retrieved by the intersection process generated by the three activation sources. In other words, an arithmetic fact is being retrieved from an organized memory representation by means of spreading activation, rather than by some numerical computation process. As one illustration, if answers (C-nodes) are weighted as heavily as parent nodes in the search, the simulation incorrectly retrieves the highest node as the correct answer; that is, if the 8 in 4 + 3 = 8 is accorded equal importance, then the C-node 8 at 4 + 4 receives the greatest amount of activation. In this situation, the simulation eventually "decides" that 8 is the correct answer to the problem. Second, the node retreived by the search process propogates activation through the network according to the embedded relationships among the facts, the "next" and "other operation" relations. The important consequence of this is that other arithmetic facts which are related to the retrieved fact have been activated by the search. This generates a reasonable and precise basis for the important decision stage effects to be described and simulated below.

Decision Stage Processes

The decision stage of the simulation receives from the search component the retrieved answer x from the network of stored facts. The task of the decision stage is to determine if this retrieved answer matches the answer stated in the stimulus problem; i.e., for a problem a + b = c, the decision stage evaluates the equation x=c? At the end of the decision stage, a judgment is reached as to the correctness of the stimulus problem, and this decision, true or false, is passed on to the response stage. An early version of the simulation revealed the need for two separate decision mechanisms, one functioning on the basis of activation levels generated during search, and one functioning on the basis of a numerical inequality judgment procedure. Both are reported here, along with an argument that both processes are characteristic of each decision making event in the empirical research. Conflicts over the different RT predictions represent a major set of new hypotheses generated by the simulation model.

Decision Process I — Symbolic Distance/Split Mechanism

In this first decision mechanism, the simulation attempts to match the answer retrieved from memory, x, to that presented in the stimulus, c. The matching process is viewed as comparable to that in the number comparison literature (e.g., Banks, 1977). That is, x and c are both accessed on some mental representation of numerical magnitude, and their identity (in the case of a match) is noticed. In Bank's terminology, x and c are located on the mental number line, and are found to occupy the same location. Under a match condition, the process is assumed to require only some constant amount of time, independent of the numerical values used.

Under mismatch conditions, this decision process will have accessed two separate number line locations, one each for x and c, and the resultant "mismatch" decision is the product of a discrimination between the two locations. As the number comparison literature indicates, such discriminations become less difficult, and less time consuming, as the two locations become more separated on the number line (recall from above that we found a 100 msec advantage for problems with splits of 5 or 6, vs. 1 or 2, in our first study of mental addition). According to Banks (1977), this discrimination process operates on a subjective number line which is compressed at the larger end of the scale -- that is, larger numbers occupy progressively "closer" locations. Thus, discriminations between pairs of locations are not simply a function of split. Instead, the compression makes discriminations between locations increasingly difficult as the numbers increase in size. Stated simply, it is harder to discriminate between 8 and 9 than between 1 and 2, despite the adjacency of the numbers, since 8 and 9 are closer together on the number line. Such a scheme inplies a relative or proportional effect of split -- the numerical difference between numbers is weighted by the size of the numbers themselves.

Stazyk et al. found exactly this sort of proportional split effect in their study of multiplication. Accordingly, the process implemented here is a proportional one, that is one in which differences between correct and stated answers are weighted by the size of the correct answer; in the simulation, this effect is approximated by the formula $1 - ((x-c)/x)$. Thus, an addition problem wrong by only 1 experiences a rather large decision stage adjustment due to the difficulty of the discrimination, and the adjustment is magnified in larger problems. For example, the problem $4 + 5 = 8$ generates a proportional split $1-(1/9) = 8/9$. This proportion is magnified for larger problems; for $9 + 7 = 17$, it is $1 - (1/16) = 15/16$. Alternately, a problem wrong by 8, like $4 + 5 = 1$, has only a small correction due to the simplicity of the discrimination, indicated by the proportion $1- (8/9) = 1/9$. Its larger counterpart, $9 + 7 = 8$, generates the proportion $1 - (8/16) = 1/2$.

Notice that the operative factor in this sort of scheme is one of difficulty in decision making. That is, the effect of the mental comparison of $\underline{x}$ and $\underline{c}$, when they do not match, is to make a decision more difficult to reach, by virtue of the more difficult discrimination required. Small values of split, in particular, are instrumental in this slowing down of the decision mechanism, in that they are so near the correct value that the discrimination is difficult. Whereas factors such as repetition or priming exert a facilitative effect on search processes, decision processes seem more prone to effects in the opposite direction. This orientation is represented here by rendering tne decision making process more difficult to complete (and by adding time to the estimated RT for that stage).

Decision Process II — Activation Mechanism

As in the previous decision process, the activation mechanism process receives the correct answer to the stimulus problem from the search stage. Unlike the previous decision mechanism, the process based on activation relies only on the search-generated amounts of activation for judging a problem as correct or incorrect. In brief, this mechanism compares the activations for the node *x*, the retrieved answer, and the most highly activated C-node, the most likely intersection node whose answer value equals *c*. Decision stage difficulty, and predicted RT, is a function of differences in the levels of activation for these two nodes.

Consider the operation of this decision mechanism when the stimulus problem is true. At the end of search, the retrieved answer *x* is passed to the mechanism for evaluation. The amount of activation associated with this *x* node, termed x_a, is an indicator of how strongly "suggested" the answer value is for the stimulus. The decision process consists of a comparison of this *x* activation level with the highest activation level achieved by a C-node, the latter being the most strongly suggested node where the answer equals the stated answer in the stimulus. Under the true or match condition, the most highly activated C-node is the same as the *x* node passed from the search stage. Since the two activations being compared are in fact the same, the decision mechanism has evidence that the retrieved node *x* is the same as *c*. Consequently, the *x=c*? equation is evaluated as true, since $x_a = c_a$ is true, and the decision mechanism judges the stimulus problem to be correct.

Decision time in this situation is a constant value.

Consider now the situation when the stimulus problem was incorrect. As always, the search stage passes the value $\underline{x}$ to decision, along with the amount of activation which has accrued at that node. Furthermore, since $\underline{c}$ was the answer indicated in the stimulus, the decision mechanism accesses that C-node with the highest amount of activation, this being the most likely answer node if $\underline{c}$ is truly the correct answer (a scan of the set of intersection nodes where the answer equals $\underline{c}$ identifies this most highly activated C-node). The decision process now compares the associated activation values of the two nodes $\underline{x}$ and $\underline{c}$, and finds them to be different. This is the basic evidence necessary to judge the stimulus problem incorrect. As was the case for decision process I, the amount of time necessary for this mismatch judgment is an inverse function of the difference, although for process II the relevant difference is in terms of activation levels. Specifically, decision time here is determined by the formula $1 - ((x_a - c_a)/x_a)$. As in the numerical inequality process, the more discrepant the two activation levels, the less likely that $\underline{c}$ is a "reasonable" competitor for the correct answer; stated in Pandemonium-like terms, a value for $\underline{c}_a$ which is quite different from $\underline{x}_a$ does not "demand" or shout for the close attention that a very similar activation level would.

The general consequence of this procedure is often quite similar to that found in the numerical inequality process — a C-node which differs substantially from $\underline{x}$ will have received less activation during search, and therefore yields less competition during the decision making process. For instance, when 8 + 7 = 10 is presented, $\underline{x}_a = 171$ and $\underline{c}_a = 80$

(compare with $\underline{x}_a$ = 195 and $\underline{c}_a$ = 102 for 8 + 7 = 14, a closer C-node which receives greater activation). On the other hand, certain situations in the empirical data, particularly when the confusion effect is considered, indicate the need for the both the activation mechanism as well as the simpler numerical inequality process. These situations are discussed in the next section.

In summary, two different decision processes are implemented in the present simulation, one based on the notion of comparisons along a symbolic magnitude dimension, and one based on relative levels of activation produced by the search stage. Both processes reflect the general orientation that decision making can be made more difficult by several factors, with a consequent slowing of decision times.

Behavior of the Simulation and an Elaborated Model

Before turning to the specific effects to be accounted for by the model, the effects of problem size, split, and confusions, it is important to mention a global aspect of the simulation's behavior. The model as it stands performs correctly on the stimulus problems given to it. This statement means more than simply that the simulation program executes correctly. It means that our network-based approach to arithmetic fact retrieval can be simulated in a psychologically plausible fashion, without undue or unreasonable assumptions. This is not to deny that factors other than retrieval may be important in adults' processing of simple arithmetic facts, as the elaborated model below attests. This is simply to say that no recourse to counting or computational procedures is necessary to provide an account of verification of addition and multiplication problems.

Problem Size Effect

At a somewhat more detailed level, the simulation accounts for the problem size effect very well[3]. As mentioned, the predicted search time in the model is a function of the target node accessibility, weighted by an appropriate coefficient (see Figure 2). Figure 4 illustrates the observed and predicted data points for the addition problems 0 + 0 through 9 + 9; Figure 5 shows these scores for the same problems in multiplication. The solid curves in the figures are the empirical functions found in Ashcraft and Stazyk (1981) and Stazyk et al. (1982), respectively. The most interesting and important aspects of the similarities here derive from the distance scheme employed in the simulation. Whereas the empirical functions on the graphs are predicted regression lines based on the structural variables sum squared and correct product, respectively, the predicted data points are derived solely from the normatively-obtained difficulty ratings (Ashcraft & Hamann, Note 4). Clearly, the majority of the predicted data points fall close to the empirical regression functions, indicating the appropriateness of the normative ratings as indices of search distance through the network representation. It is especially interesting that the RT predictions based on rated difficulty yield the exponential pattern of RTs obtained in the empirical studies, since the exponential shape has been taken as evidence against computation mechanisms like counting (see Ashcraft & Battaglia, 1978). Just as clearly, there are systematic departures from the empirical functions in the predicted data points, departures which necessitate the elaboration described later.

Figures 4 and 5 about here

Split Effect

Recall that Decision Process I simulates a symbolic distance effect in the decision stage. That is, decision difficulty and time are inverse functions of split, the difference between the correct and stated answers weighted by the absolute magnitude of the correct answer (see Figure 2). The reasoning behind this process is that numbers which are close on the symbolic magnitude continuum are difficult to discriminate, generating slow decision times. Figure 6 illustrates the performance of the simulation on the set of addition and multiplication problems formed with the digits 4 and 5, with answers varying from a split of 1 to a split of 12. Compared to the predicted RT of 1080 msec for the true 4 + 5 = 9, a split of 1 (4 + 5 = 8 or 10) generates a 133 msec slowing of RT under decision process I (solid functions). The amount of slowing, indicating difficulty in decision making, decreases as split increases, as has been found repeatedly in our empirical studies (see especially Ashcraft & Stazyk, 1981). One minor feature of note is that this split effect is generated on a proportional basis here, whereas our studies of addition have examined split only on an absolute, non-proportional basis. It would appear that the Stazyk et al. suggestion of a proportional scheme for both addition and multiplication was a reasonable interpretation of their data. Figure 7 shows comparable predictions for the addition and multiplication problems formed with the digits 8 and 7.

Figures 5 and 6 about here

As the figures show, Decision Process II (dotted functions) generates RT predictions in addition which are often quite similar to those from model I, although the functions are rather irregular in shape. The source of the irregularity in process II's predictions is accrued activation at C-nodes, of course. Thus for 8 + 7, the absence of the answers 22 and 23 in either network means no activation will have accrued to these C-nodes, leading to the prediction of rapid RT due to lack of competition. The elevated prediction at c=4 for the problem 4 + 5 (compared to c=3) is due to the presence of a C-node equal to 4 on the family of answers "4 + b"; this C-node not only receives activation as a family member, it also transmits activation back towards the x node (notice that an answer equal to 3 is not in either set of family members). It is not known whether the smoothly decreasing split function or the more irregular activation function is more accurate under the addition operation, since no studies are available which have varied split in such a systematic fashion. It is entirely possible that the standard procedure of averaging RTs across problems, with split held constant, has obscured these irregular patterns. These possibilities have become apparent only through the development of the present model.

On the other hand, Process II does make several incorrect predictions under the multiplication operation when the simple split effect is considered. For instance, the problem 4 X 5 = 22 is predicted by process II to be quite rapidly rejected, since no C-node competition exists. As Stazyk et al., showed, however, there is an effect of split in multiplication which is independent of the "legality" or presence of answers in the table. This failure is specific to multiplication, since only some of the numbers in the range 0 to 100 occur as answers in the multiplication table. In other words, while a problem like 8 X 7 = 52 is rather difficult for subjects to reject (see Experiment I in Stazyk et al.), the activation-based decision process experiences no difficulty in rejecting the stated answer of 52. The simple reason for this is that 52 is not an available value of c in the multiplication (or addition) table, so no activation accrues to any C-node in either table. This is in fact a general aspect of the simulation's performance to multiplication problems; only those incorrect problems with legal multiplication answers experience decision stage difficulty due to spreading activation. Interestingly, values of c which are legal but unrelated to the stimulus (e.g., 4 X 5 = 21) do generate slower RT predictions than those for illegal values of c (4 X 5 = 23), an effect termed the "privileged answer" effect by Duffy and Fisher (Note 2).

Confusion Effect

Despite the incorrect predictions generated by the activation process under simple split conditions, the figures also illustrate the success of this process, and the relative failure of the simpler split model, under conditions of "confusions". To repeat, stimulus problems given a value for c which is correct under another operation, such as 4 + 5 = 20, are found to require a great deal of extra processing time in the empirical research (Winkelman & Schmidt, 1974). Furthermore, the confusion effect emerges quite strongly even without this "cross-operation" manipulation; Stazyk et al. found that values of c which are multiples of either a or b (for example, 6 X 7 = 36) yielded up to a 300 msec slowing of RT.

As the figures illustrate, both of these confusion types are predicted by the activation mechanism II. For the 8 X 7 problem set, for instance, each stated answer which is a multiple of either 8 or 7 (e.g., 32, 40, 48) shows a slowing of predicted RT, contrasted with the split process predictions. Furthermore, the cross-operation confusions are invariably predicted to be quite lengthy, especially in addition (e.g., 4 + 5 = 20). Decision process I fails to predict any special slowing for such problems. The reason for this is quite clear -- for 4 + 5 = 20, the answer is wrong by 11, thus the proportional adjustment (in a sense an index of how much difficulty is generated by the split) is 1 - ((9 - 20)/9) = -2/9 (since split is bi-directional, the absolute value of x-c is always taken). Stated simply, decision process I judges 20 to be too far from 9 to produce much decision difficulty. Obviously, such a simple mechanism misses an important aspect of the stimulus, that the numbers 4, 5, and 20 are highly related under the multiplication operation.

It is exactly this "other operation" relation which is captured by the networks simulated here, and by the parallel activation of the multiplication network during search. This parallel activation is functional in a problem like 4 + 5 = 20 since the activation of the C-node 20 combines with the parent node activation in the multiplication table. Thus, the 4 X 5 = 20 node in the multiplication network becomes the strongest competitor during decision due to the "other operation" relation from the 4 + 5 = 9 node, and to the activation of all C-nodes with answers equal to 20. Furthermore, the accurate predictions of the confusion effect within the same arithmetic operation reflect the other important network relation captured by the simulation, that successor nodes along families are interrelated (in the "semantic relatedness" sense). Thus, Kintsch's (1974) generalization, that semantic relatedness slows negative decisions, applies fully to the present model, and to mental arithmetic performance in general.

Pragmatically, the longer of the two RT predictions at any problem seems to model the empirical effects. While this scheme "works", it lacks a satisfactory theoretical rationale. One way to develop such a rationale would be to consider the decision stage subject to various factors which retard the decision making process, as suggested above. On this view, two factors which increase the difficulty of decision making have been simulated here, nearness of a stated and retrieved answer on the mental number line, which implies difficulty in discrimination, and similarity of stated and retreived answers from the standpoint of search-generated activation levels, again with an implied difficulty in conflict resolution or decision making. If the decision stage is conceptualized as one which is easily disrupted by competing information, then both processes may be manifestations of the same underlying activity. A second possibility for unifying these effects[4] also exists. We may reconceptualize the symbolic distance effect (Banks, 1977) as an effect of spreading activation along the mental continuum of number magnitude itself. On such a view, the value x sent to the decision stage activates its representation on the subjective number line, and propogates a gradient of activation to its neighbors. Clearly, when split is small, the location on the number line which corresponds to c is likely to have received some of this spread of activation, yielding the sort of competition that process II involves. Note that such a scheme will even predict the slowness of an illegal multiplication answer like 41 for 7 X 6 — since 41 is quite near 42 on the number line, the decision process will require more time to perform the relevant discrimination. Such a spreading activation process along the mental number line would of course be joined with search-generated activation levels in decision making. That is, the lesser activation of 36 from the x value 42 along the number line (in the problem 7 X 6) would be modified by the amount of activation accumulated during search at the C-node 36 in the network. Thus this reinterpretation still requires a number line (process I) and an activation-based (process II) comparison for decision making. The advantage, however, would be that both processes stem from the same overall notion of spreading activation through mental representations of number.

Lack of Fit and an Elaborated Model

The most disturbing lack of fit between predicted and observed data in the simulation (see Figures 4 and 5) occurs for those problems in addition and multiplication which involve a or b values of zero or one.[5] The simulation generally predicts RTs which are too fast for these problems. The most clear-cut example of this is the set of multiplication problems involving a or b values of zero. As Stazyk et al. (see also Parkman, 1972) found, these problems were quite slow to verify, with RTs ranging from 1100 to 1750 msec., and quite prone to error. And yet, in the simulated data, these problems uniformly fall below the regression curve, with very rapid predicted RTs.

The reason for this discrepancy involves the subjectively derived difficulty measures which are used for network distance estimates. Despite the fact that they perform quite slowly to problems like 7 X 0 = 0, subjects rate these problems as quite easy -- on the original scale from 0 to 9, 7 X 0 = 0 has a mean difficulty rating of .42. Our suspicion, noted in Stazyk et al., is that subjects' ratings for these problems derived from their knowledge of the rule that "anything times zero is zero", a simple rule suggesting low difficulty. We thus confront the question of how many of the simple facts in addition and multiplication were rated via such rules, and how seriously the ratings should be taken.

Several investigators (see Baroody, in press, for example) have suggested a rule-based "procedural knowledge" basis for performance to these simple problems, and the model in Figure 1 specifically included such rule-based heuristics and algorithms as part of mature subjects' knowledge. It seems very appropriate, then, to conclude that the lack of fit illustrated in Figures 4 and 5 is largely due to the absence of such rule-based performance in the simulation. Specifically, the simulation now retrieves the answer 0 to the problem 7 X 0, with search distance estimated by the rated difficulty value. If this difficulty value is in fact derived from rule knowledge, then it should be a poor estimate of search distance per se. It seems that a more adequate model should isolate problems like the zero facts in multiplication as "special cases", to be processed via the procedural knowledge component rather than the usual network retrieval process.

I have elaborated the present simulation model to incorporate this notion of procedural-based performance for four notable "special cases" in addition and multiplication, addition or multiplication by zero or one. The elaboration involves two important elements -- first, a reasonable method by which the need for rule-based processing can be determined, and second a set of procedures which specify the rules used for such processing. The first element, serving to "call" the procedural knowledge routines, is handled as follows. It seems clear that problems like the zero facts in multiplication are not nearly as easy as subjects' ratings would suggest. In other words, if we rely on ratings like 1.54 for the problem 5 X 5, we cannot take seriously the .42 for 7 X 0. Accordingly, all difficulty ratings for the special cases where either a or b equals 0 or 1 have been altered; in the accessibility metric used here, these problems have been reassigned the arbitrary value of 30. The processing consequence of this is that any such problem will fail to achieve the activation criterion of 100 during search; in psychological terms, these problems have such low accessibility levels that they cannot be retrieved from the network. Upon such retrieval failure, the procedural knowledge component of the simulation is re-entered, and the set of rules for the special cases is accessed[6].

I include only a set of three "procedures" in the current elaborated model, since no especially compelling evidence exists to suggest other more sophisticated procedures (but see for example Groen & Poll, 1974, or Resnick & Ford, 1982, for such procedures in subtraction). The three components are (1) a general counting procedure, (2) a rule for multiplication by zero, and (3) a rule for multiplication by 1. By a matching process, in which the values for a and b as well as the sign (operator) are considered, the procedural component selects one of these three procedures to use in solving the problem. Thus, the presence of the multiplication sign eliminates the first component from consideraton, and a multiplication sign and a 1 eliminates the second component as well. The presence of the addition sign invariably calls the counting rule, regardless of the a and b values. No reasonable basis for estimating the time parameters for the two multiplication rules is available at present, although I suspect that rule 3 will generate faster responding than rule 2. On the other hand, the counting algorithm in component 1 functions at the rate of 200 msec. per increment, this on the evidence contained in Ashcraft and Bartolotta's (Note 5) experiment. In all cases, the relevant procedural knowledge is accessed, and a correct answer x is generated. This value is passed to a skeleton decision stage for the x=c? comparison (the full set of decision stage processes, especially those involving activation levels, cannot apply since network activation was abandoned midstream).

While it is somewhat unfortunate that such a post hoc revision of the network distances and the internal processes of the simulation is necessary, it is probably desirable from a psychological standpoint that such elaborations be explicitly included in the model. That is, even though most arithmetic performance on the simple facts of addition and multiplication may be accomplished via fact retrieval, subjects are perfectly capable of generating a rule-based explanation for their performance (Ashcraft & Hamann, Note 6; for instance, justifying 11 as the sum of 8 + 3 by saying "You can count up three like 9, 10, 11, and you have it."). Furthermore, in the discussion below, it is argued that the development of arithmetic performance in children is largely a shift from various rule-based procedures like counting to search processes as simulated here. It would be unreasonable to assert that rules used by children are totally lost from the knowledge base simply because a new basis for performance is acquired later.

Discussion

This final section begins by considering several relatively simple modifications of the simulation which would extend its generality to different tasks and experimental situations. This is followed by a more substantive consideration of developmental issues in arithmetic performance, and the sorts of elaborations necessary for the present simulation.

Discussion

This final section begins by considering several relatively simple modifications of the simulation which would extend its generality to different tasks and experimental situations. This is followed by a more substantive consideration of developmental issues in arithmetic performance, and the sorts of elaborations necessary for the present simulation to provide a useful account of the development of arithmetic knowledge and performance.

To begin with, the original purpose of the simulation was to model network retrieval and decision processes in adults (Ashcraft, 1982; Stazyk et al., 1982). Not surprisingly, various methodological and design characteristics of the original experiments appear in the simulation, at least by default. To take just two examples, our experiments have always tested subjects in only one of the two arithmetic operations at a time, either addition or multiplication. There is no evidence on the sorts of modifications that might be necessary to simulate performance on mixed trial procedures. Secondly, we have carefully avoided stimulus sequences in which there is any repetition of numbers from one trial to the next. While there are indications that repetition effects will appear in RT measures (see Ashcraft & Battaglia, 1978, for instance), these are outside the scope of the present model.

Somewhat more generally, however, the simulation can be easily extended to make useful predictions in different kinds of tasks. For instance, Ashcraft and Stazyk's Experiment 2 tested subjects on addition problems with sums up to 30, rather than merely on the basic addition facts with sums up to 18 (we had stimuli such as 14 + 13 = 27 and 16 + 9 = 25). No particularly major overhaul of the simulation is necessary in order for it to process these kinds of problems, especially since the problem size effect found for these larger problems is well predicted by distance estimates to the simple combinations addressed here. Nor would the simulation require major reworking to yield predictions on a task which asked if <u>any</u> simple arithmetic relationship obtains for a set of numbers ("yes for '5 6 11', "no for '8 7 41'). Clearly, the model would predict that time to make a "yes" decision would be a function of node accessibility, and time for "no" decisions a function of split and confusion factors. No priority of one operation over another would be predicted, the decision being governed instead by the highest accessibility (shortest time) for a relevant intersection. Such a task might also reveal more clearly the role of the hypothesized "global evaluation" process (Ashcraft & Stazyk, 1981) in arithmetic knowledge and performance.

Finally, our experiments have never inquired systematically into the possible effects of residual activation on later performance, although both facilitation and inhibition effects across trials have been demonstrated in other areas (for instance, semantic relatedness effects from trial _n_ to trial _n+1_ have been demonstrated by Ashcraft, 1976; Haviland & Clark, 1974; and Loftus & Loftus, 1974). The simulation now behaves as if each trial were unrelated to its predecessor, although only a minor modification is required to prevent the networks from reverting to inactivation prior to a new trial. This modification would simulate an hypothesized decay of activation process (see Collins & Loftus, 1975, for example), but would require further empirical work to determine suitable decay parameters.

At a much broader level, more extensive modifications to the simulation would be required to simulate the effects obtained in developmental investigations of arithmetic performance. Ashcraft (1982) has presented a table showing the developmental progression of the problem size effect from grade 1 through college. The major change, aside from a decrease in the intercept estimates, is a flattening of the problem size effect. In other words, as children grow older and more experienced, the problem size effect becomes attenuated (although it of course remains highly significant). I have suggested that part of this speeding is a consolidation of the fact retrieval process, particularly in elementary grades (Ashcraft & Fierman, 1982). Note that consolidation is essentially an issue of overlearning or mastery, and would be represented appropriately by strengthening initially low accessibility values in the networks. Yet another part of the remaining developmental change would appear to reflect increasing automaticity (see also Resnick & Ford, 1982), such that children's retrieval from about the sixth grade level becomes less conscious and more automatic. So far, this prediction has not been tested. It is unclear how an increase in automaticity could be represented in the model (except by the trivial method of revising the coefficients which govern the RT predictions).

More important than these changes, however, is the issue of children's knowledge about arithmetic, the "procedural" component of Figure 1. There is extensive evidence, certainly for children up through about third grade, that retrieval is not the only, or even the major, process by which simple addition is performed. As Groen and Parkman (1972) found, first graders' performance is best predicted by a counting variable (also Gelman & Gallistel, 1978; Ginsburg, 1977; Siegler, Note 7). Indeed, many such informal procedures appear to be available to children as they perform the simple arithmetic problems. This assortment of procedural knowledge is acknowledged in Figure 1, but is represented in the simulation only by the minor rules for performance to "special cases". For children, there are probably many other classes of problems which are routinely performed via procedural knowledge, and at early stages of arithmetic instruction only a very few problems stored in the networks (see Baroody, in press). A viable developmental model of performance must reflect this rich variety of performance bases in children.

The elaborated model presented here can serve as a guide to such a developmental model of arithmetic performance. Obviously, the simple counting algorithm used here for "special cases" could be applied to an addition problem, with predictable consequences for search time and network activation levels. The evidence suggests that this surely is a commonly used element of children's processing. More generally, by carefully analyzing children's performance to simple arithmetic problems, a comparatively small number of general purpose rules (heuristics and algorithms in Figure 1) may be identified, and then included in the simulation. Each of these would include not only the action to be taken as a consequence of accessing the rule (counting up by ones, for example), but also some index of the strength of the rule. Attention to such a "precedence" scheme among rules would not only dictate which of several possibly relevant rules would be selected for a particular problem (see Anderson, 1982), but also would then provide a more meaningful way of approaching the proposed developmental shift from rule-based performance to fact retrieval. Note that the current scheme, retrieval failure followed by rule-based performance, achieves a degree of support from Siegler's (Note 7) study of pre-schoolers; his data indicated that any overt strategy for addition problems required more time than basic fact retrieval for these children, and also was accompanied by one or more indications of literal counting (use of fingers, for example). On the other hand, there may be cases in which a procedure has been "overlearned" (e.g., to add 9, first add 10 then count back 1) to a degree that a stimulus like 6 + 9 might trigger the special rule first, bypassing the normal precedence of retrieval over procedural solution.

At this point, it would be entirely speculative to assert that attempted retrieval <u>always</u> occurs before rule-based attempts at solution. It is certainly conceivable that some "overlearned" rules are accessed first if their precedence values exceed that of the fact retrieval procedure. The issue of precedence among the various components of procedural knowledge, including differential precedence for the fact retrieval process, counting, rules for special cases, and so forth, may indeed be a very promising one for models of developmental change in arithmetic processing.

Footnotes

1. Fact retrieval is an elementary procedure, as conceived here, which directs the search attempt through the declarative knowledge store. Since its involvement in every fact retrieval operation is a constant component of the search, the retrieval procedure per se does not add to the search time in a way which can be separated from the network search time. Unless noted otherwise, the term "procedural knowledge" in this paper follows the usage suggested in Figure 1 -- that is, the term is reserved for non-retrieval procedures such as counting, estimating, or solving by a rule. Thus the terms declarative knowledge and procedural knowledge correspond, respectively, to "knowledge of arithmetic facts" and "knowledge about arithmetic".

2. Why shouldn't all nodes which have received activation, all family members and C-nodes alike, propogate a spread of activation to their neighbors, instead of the current scheme of all C-nodes but only the highest node of intersection from the set of family members? The two reasons for this entail economy of program execution and net efffect on processing. When the first parent node activates its family, all 10 columns in each table contain an active element; likewise for all 10 rows when the second parent activates its family. Thus, each of the 100 elements in a table is on both a column and a row which would receive activation from a family member. An initial version of the simulation propogated activation from each family member, in the usual decreasing gradient fashion. This scheme required excessive amounts of execution time for the

program, since each of the 200 "grandchild" nodes (parent node to family member, family member to its family members) was being updated once as a grandchild of parent #1 and a second time as a grandchild of parent #2 (and then often as a neighbor of the C-nodes as well). Such massive activation spreading, furthermore, had no functional effect on later processing; the entire network was boosted in activation with this exhaustive spread, but the relative levels of activation between any pair of nodes were not appreciably different from those found in the simpler scheme now used. Since the extra computational effort yielded no functional differences, the more complete activation spreading procedure was dropped.

3. Absolute values of predicted RT should not be taken as seriously as relative RT effects, such as the problem size or confusion effect, which model the empirical RT differences between pairs of stimuli or experimental conditions. Such relative comparisons are appropriate for any orthogonal contrasts from the model, since the same internally consistent metric has been used throughout. Note also that the RT predictions are in a sense estimates of population effects — that is, no error variance is added to the predicted scores. Accordingly, no tests of significance between observed and predicted data are reported.

4. I thank Jim Greeno for this insight.

5. Predictions for tie problems (the circled points in Figures 4 and 5) are somewhat slow in the simulation, although as a class they are faster than non-tie problems. The remaining facilitation in observed RTs may be due to the perceptual ease of encoding repeated digits for a and b.

6. Miller, Perlmutter, and Keating (in press) tested adults in a production task, and found the "zero" multiplication facts to be quite rapid. They suggest that the slowness of these problems in Parkman's (1972) and Stazyk et al.'s (1982) data is an artifact of the verification task. It is not at all clear how such a reliable artifact might arise, or how it should be modeled. If accessibility values for these problems are indeed high, then some decisional uncertainty would be implicated in the verification task. Alternately, a procedural solution for these problems might occur prior to a retrieval attempt, with the same sort of decision stage uncertainty slowing down verification RT. In any event, the uncertainty of the decisions, contrasted with the other multiplication facts, surely suggests that something is unusual for these problems.

Reference Notes

1. Ashcraft, M.H. & Bartolotta, R. Adding and Counting in the First Three Years of Grade School. Manuscript in preparation.

2. Duffy, S.A. & Fisher, D.L. The organization and processing of multiplication facts. Paper presented at the Midwestern Psychological Association, St. Louis, May, 1980.

3. Bartolotta, R. A Comparison of Cross- and Within-Operation Confusion Effects in Two Adult Populations. M.A. Thesis, in progress.

4. Ashcraft, M.H. & Hamann, M.S. Normative Difficulty Ratings on the Basic Arithmetic Facts. Manuscript in preparation.

5. Ashcraft, M.H. & Bartolotta, R. Mental Counting. Paper presented at the Psychonomic Society, Minneapolis, November, 1982.

6. Ashcraft, M.H. & Hamann, M.S. Children's Strategies for Solving Simple and Complex Addition Problems. Paper presented at the Psychonomic Society, Philadelphia, November, 1981.

7. Siegler, R.S. Strategy Choice in Addition: Children Have Been Right All Along. Unpublished manuscript (Department of Psychology, Carnegie-Mellon University, 1982).

References

Anderson, J.R. A spreading activation theory of memory. Journal of Verbal Learning and Verbal Behavior, 1983, 22, 261-295.

Anderson, J.R. Acquisition of cognitive skill. Psychological Review, 1982, 89, 369-406.

Ashcraft, M.H. Priming and property dominance effects in semantic memory. Memory & Cognition, 1976, 4, 490-500.

Ashcraft, M.H. The development of mental arithmetic: A chronometric approach. Developmental Review, 1982, 2, 213-236.

Ashcraft, M.H. Procedural knowledge versus fact retrieval in mental arithmetic: A reply to Baroody. Developmental Review, in press.

Ashcraft, M.H. & Battaglia, J. Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 527-538.

Ashcraft, M.H. & Fierman, B.A. Mental addition in third, fourth, and sixth graders. Journal of Experimental Child Psychology, 1982, 33, 216-234.

Ashcraft, M.H. & Stazyk, E.H. Mental addition: A test of three verification models. Memory & Cognition, 1981, 9, 185-196.

Banks, W.P. Encoding and processing of symbolic information in comparative judgments. In G.H. Bower (Ed.), The Psychology of Learning and Motivation (Vol. 11). New York: Academic Press, 1977.

Banks, W.P., Fujii, M., & Kayra-Stuart, F. Semantic congruity effects in comparative judgments of digits. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 435-447.

Baroody, A.J. The development of procedural knowledge: An alternative explanation for chronometric trends of mental arithmetic.

Developmental Review, 1983, in press.

Collins, A.M. & Loftus, E.F. A spreading-activation theory of semantic processing. Psychological Review, 1975, 82, 407-428.

Friedman, A. Memorial comparisons without the "mind's eye." Journal of Verbal Learning and Verbal Behavior, 1978, 17, 427-444.

Gelman, R. & Gallistel, C.R. The child's understanding of number. Cambridge, Mass.: Harvard Univ. Press, 1978.

Glass, A.L., Holyoak, K.J., & O'Dell, C. Production frequency and the verification of quantified statements. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 237-254.

Ginsburg, H. Children's arithmetic: The learning process. New York: Van Nostrand, 1977.

Groen, G.J. & Parkman, J.M. A chronometric analysis of simple addition. Psychological Review, 1972, 79, 329-343.

Groen, G.J. & Poll, M. Subtraction and the solution of open sentence problems. Journal of Experimental Child Psychology, 1973, 16, 292-302.

Haviland, S.E. & Clark, H.H. What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 237-254.

Holyoak, K.J. & Walker, J.H. Subjective magnitude information in semantic orderings. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 287-299.

Kintsch, W. The representation of meaning in memory. Hillsdale, N.J.: Erlbaum, 1974.

Loftus, G.R. & Loftus, E.F. The influence of one memory retrievalon a subsequent memory retrieval. Memory & Cognition, 1974, 2, 467-471.

Loftus, E.F. & Suppes, P. Structural variables that determine the speed of retrieving words from long term memory. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 770-777.

Miller, K., Perlmutter, M., & Keating, D. Cognitive arithmetic: Comparison of operations. Journal of Experimental Psychology: Learning, Memory, and Cognition, in press.

Moyer, R.S. & Landauer, T.K. The time required for judgments of numerical inequality. Nature, 1967, 215, 1519-1520.

Parkman, J.M. Temporal aspects of simple multiplication and comparison. Journal of Experimental Pshchology, 1972, 95, 437-444.

Resnick, L.B. & Ford, W.W. The psychology of mathematics for instruction. Hillsdale, N.J.: Erlbaum, 1982.

Stazyk, E.H., Ashcraft, M.H., & Hamann, M.S. A network approach to mental multiplication. Journal of Experimental Psychology: Learning, Memory, and Cognition, 1982, 8, 320-335.

Svenson, O. Analysis of time required by children for simple additions. Acta Psychologica, 1975, 39, 289-302.

Whaley, C.P. Word-nonword classification time. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 143-154.

Winkelman, H.J. & Schmidt, J. Associative confusions in mental arithmetic. Journal of Experimental Psychology, 1974, 102, 734-736.

Figure Captions

1. Information processing model of mental arithmetic knowledge and performance (from Ashcraft, 1982).

2. Summary of processing in the simulation model, with predicted RT equations.

3. Network representation of addition and multiplication tables for the problem 4 + 3 = 7 and 4 X 3 = 7. The "next" relationship points 'north, south, east, and west' from each node; since all adjacent nodes are so interconnected, the "next" pointers are not displayed. Similarly, an "other operation" pointer connects each node to its corresponding location in the alternate table.

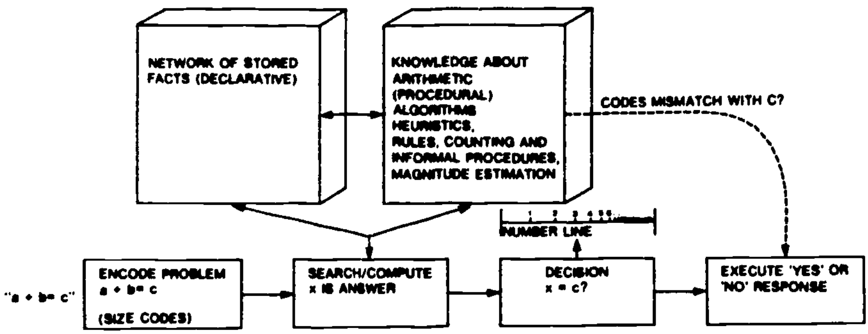
4. Observed and predicted RT scatterplots for the problem size effect in addition. Circled points are tie problems (e.g., 4 + 4 = 8); points flanked with dashes are those with a zero addend (e.g., 7 + 0 = 7). The solid function is the best fitting regression line from Ashcraft & Stazyk (1981), which used correct sum squared as the predictor variable.

5. Observed and predicted RT scatterplots for the problem size effect in multiplication. Circled points are tie problems (e.g., 4 X 4 = 16); points flanked by dashes are those with a zero multiplier (e.g., 7 X 0 = 0). The solid function is the best fitting regression line from Stazyk et al. (1982), which used correct product as the predictor variable.

6. Predicted RTs for the set of problems formed with the digits 4 and 5 (e.g., 4 + 5 = c, 4 X 5 = c). The values of c simulated on the two halves were incremented by one's.

7. Predicted RTs for the set of problems formed with the digits 8 and 7 (e.g., 8 + 7 = c, 8 X 7 = c). The values of c simulated for the

addition problems were incremented by ones's; under multiplication, values of $\underline{c}$ were incremented by four's, so interpolation in the multiplication function under process II is not accurate.



$RT = t_{ENCODE} + t_{SEARCH/COMPUTE} + t_{DECISION} + t_{RESPONSE}$

WHERE $t_{S/C}$ = f (NETWORK DISTANCE AND PROCEDURAL INVOLVEMENT)

AND t_D = NEGATIVE f(SPLIT)

AND $t_E + t_R$ = K(y - INTERCEPT)

Figure 1.

ENCODE:

$$t_e = k_e$$

PROCEDURAL KNOWLEDGE:

Call Search

SEARCH:

$$t_s = k_s x_a$$

DECISION:

	I	II
	$t_{dI} = k_p D_I + k_d$	$t_{dII} = k_p D_{II} + k_d$
where	$D_I = 1 - ((x-c)/x)$	$D_{II} = 1 - ((x_a - c_a)/x_a)$

RESPONSE:

$$t_r = k_r$$

$$RT = t_e + t_s + t_d + t_r$$

True $\quad RT = \underline{K} + k_s x_a \qquad$ where $\underline{K} = k_e + k_d + k_r$

False-I $\quad RT = \underline{K} + k_s x_a + k_p D_I$

False-II $\quad RT = \underline{K} + k_s x_a + k_p D_{II}$

Legend: Time to encode the stimulus, t_e, is a constant value k_e.

Time for search is a search constant, k_s, times the accessibility value of the node of intersection, $\underline{x}_a$.

Time to decide "true" under both models is a constant, k_d. Under Process I, time to decide "false" is the same decision constant plus time to make the numerical magnitude comparison, a difference value D_I times the proportion coefficient k_p. Under Process II, time to decide "false" is the decision constant k_d plus the difference value D_{II} times the proportional coefficient. For Process I, the difference value depends on $\underline{x}$ and $\underline{c}$, the correct and stated answers, respectively. For Process II, the difference values depends on the levels of activation of these two nodes.

Time to respond is a constant value, k_r.

Figure 2

	0	1	2	3	4	5	6	7	8	9
0				(3,94)				(7,93)		
1				(4,87)			(7,86)			
2				(5,87)		(7,82)				
3				(6,85)	(7,82)					
4	(4,91)	(5,87)	(6,84)	(7,81)	(8,85)	(9,79)	(10,76)	(11,71)	(12,75)	(13,73)
5			(7,81)	(8,79)						
6		(7,88)		(9,76)						
7	(7,93)			(10,75)						
8				(11,78)						
9				(12,75)						

(ADDITION)

	0	1	2	3	4	5	6	7	8	9
0				(0,98)						
1				(3,94)				(7,94)		
2				(6,94)						
3				(9,92)						
4	(0,95)	(4,95)	(8,93)	(12,89)	(16,91)	(20,90)	(24,88)	(28,85)	(32,84)	(36,94)
5				(15,88)						
6				(18,87)						
7		(7,94)		(21,89)						
8				(24.88)						
9				(27,87)						

(MULTIPLICATION)

Figure 3

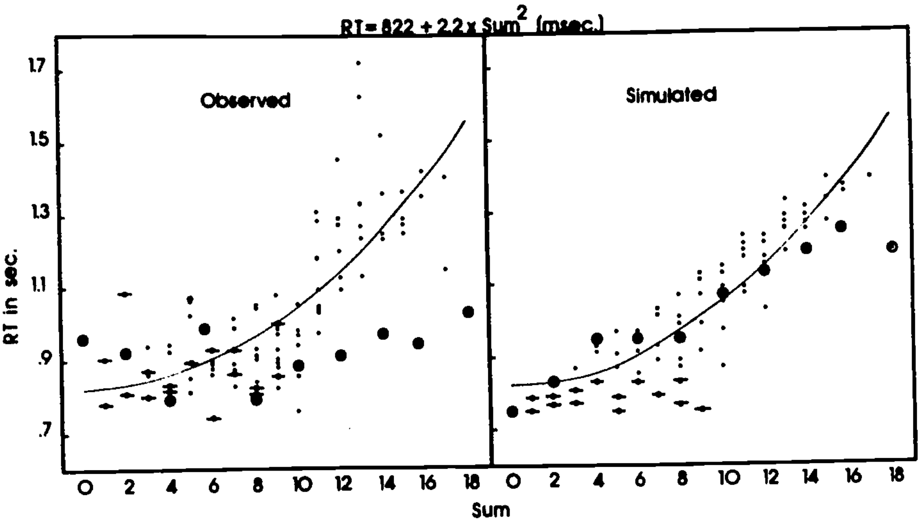


Figure 4

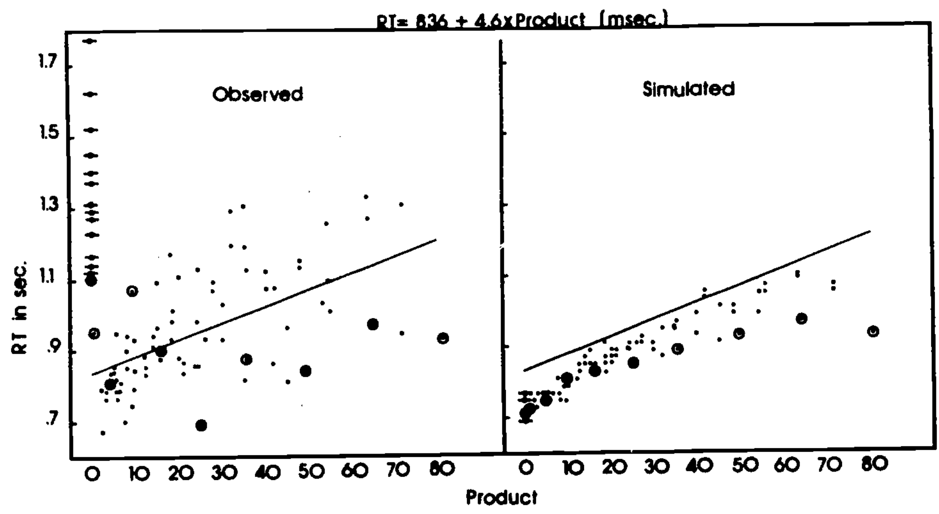


Figure 5

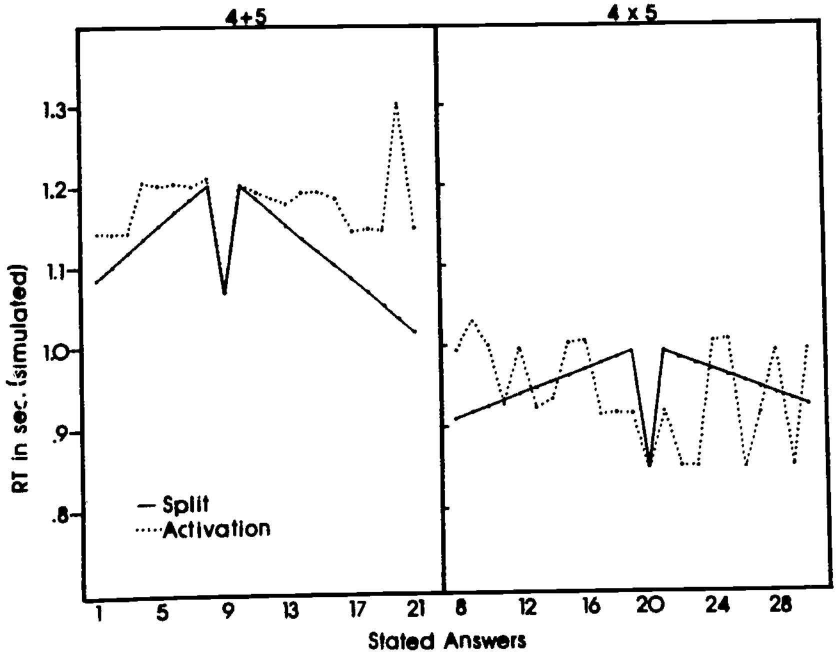


Figure 6

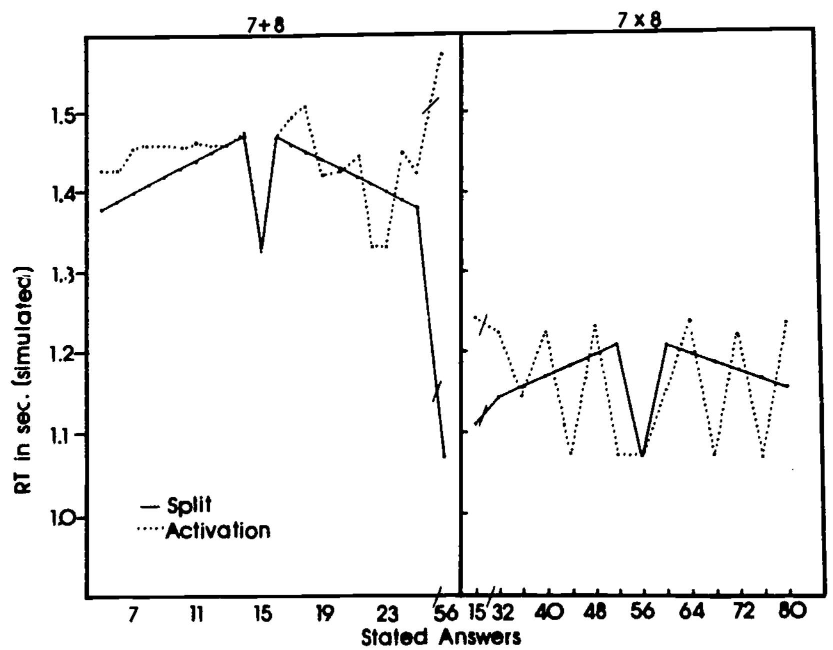


Figure 7